

多种前兆观测数据的合成信息熵研究^{*}

王海涛 杨马陵 杨又陵 曲延军 许秋龙 赵翠萍 朱 燕

(新疆维吾尔自治区地震局, 乌鲁木齐 830011)

摘要 采用概率化预处理方法, 把多种前兆观测值时间序列转化为概率值时间序列. 在此基础上, 应用信息论基本原理, 计算由多种前兆观测项目构成的信源系统的合成信息熵. 通过对乌鲁木齐及邻近地区 17 种前兆观测资料的实际计算, 发现在多次中强地震前出现半年至一年的减熵异常过程.

主题词 数据处理 地震前兆 概率化 合成信息熵 群体异常信息

1 前言

人们在实际观测、预报中发现, 不同地区不同类型地震的前兆异常图象是不尽相同的, 甚至完全不一样. 对单项前兆观测项目而言, 在不同时期不同地震前, 其异常变化可能形态各异, 无法用相对定量的标准去分析其特征. 而对一定范围内多种前兆观测资料而言, 在一次较大地震前也同样会不同程度地表现出不同形式的异常变化, 很难用一个统一的标准对前兆群体异常进行识别. 因此, 如何从多种前兆的复杂变化图象中提取出综合异常判据, 这是地震工作者一直在研究的问题.

本文在借鉴前人的大量研究的基础上^[1~3], 从信息论基本原理出发, 利用合成信息熵方法研究多种前兆群体异常信息变化, 提取地震前兆群体异常的识别标志.

2 资料预处理——分布概率的计算

信息熵的计算, 关键在于给出系统的状态概率. 我们用两种方法对前兆观测数据进行概率化计算.

2.1 极值区间法

设前兆观测值时间序列 $X_t = X_1, X_2, X_3, \dots, X_n$.

取观测值的最大值 $X_{\max}$ 和最小值 $X_{\min}$ 作为样本分布区间的上下限, 将其分为 M 个子区间, 每个子区间的宽度为 ΔX_i , 即

$$X_{\max} - X_{\min} = \sum_{i=1}^M \Delta X_i \quad (1)$$

定义序列在第 i 个子区间的分布概率

$$P_i = \frac{n_i}{N} \quad i = 1, 2, \dots, M \quad (2)$$

本文 1996 年 10 月 21 日收到.

^{*} 地震科学联合基金(95162)资助项目.

第一作者简介: 王海涛, 男, 1961 年 2 月生, 副研究员, 现从事地震分析预报研究工作.

式中 n_i 为研究时段内观测序列值位于第 i 个区间的个数; N 为研究时段内观测序列值总样本数; M 一般定为 15, 也可根据实际资料调整。

2.2 方差区间法

计算研究时段内观测值序列 $\{X_i\}$ 的均值 X , 方差 ΔX , 以 $X+k\Delta x$ 和 $X-k\Delta x$ 分别作为 (1) 式中的 $X_{\max}$ 和 $X_{\min}$, 同样原理可以求出概率序列 P_i . 在计算中, 当测值序列出现超出 $X+k\Delta x$ 上下限时, 则设定测值等于边界值. 倍方差系数 k 一般取 2.0 和 2.5 (资料较平稳) 或 3.0 和 3.5 (资料变化大)。

利用上述方法可以把前兆观测数据的整点值、日均值、旬均值及月均值等实际观测值时间序列转化为概率值时间序列。

3 合成信息熵方法

朱令人在研究地震活动图象时提出了合成信息熵方法^[1]. 我们将该方法引入对前兆观测资料的分析研究中. 研究中设一定范围内多种前兆观测资料的变化是相互独立的, 并可以构成一个信源系统^[9], 因此可以求取不同时间段内信源系统的合成信息熵。

计算中以分布概率 $P_i(x)$ 作为状态概率, 则不出现这种状态的概率为 q_i , $q_i(x) = 1 - P_i(x)$, 此时该信息元的信息熵为

$$H_i(x) = -k[P_i(x)\log P_i(x) + q_i\log q_i(x)] \quad (3)$$

由熵的可加性, 可求出前述信源系统的合成信息熵:

$$H(X) = \sum_{i=1}^n H_i(x) \quad (4)$$

为了与熵及其它指标进行比较, 需对其进行归一化, 使计算值在 (0, 1) 区间变化. 归一化合成信息熵定义为:

$$H_G = \frac{H(X)}{H_M(X)} \quad (5)$$

$H_M(X)$ 为系统极大熵值, 即取各信息元最大背景概率参与 (3) 式和 (4) 式计算. 此时合成信息熵值亦为最大。

4 资料选取与计算结果分析

为了验证合成信息熵方法的实用性, 选取了乌鲁木齐周围 100 km 范围内 17 项定点前兆观测资料 (地磁 2 项, 形变 6 项, 地电 4 项, 地下水 5 项) 进行合成信息熵的计算. 以乌鲁木齐周围 200 km 左右范围内的 5 级以上中强地震做为预测对应样本. 图 1 为前兆观测区及震中分布图. 计算之前对部分资料进行趋势项校正。

以考察时间前两年的资料计算背景概率值, 滑动式推进. 例如, 考察时间为 1993 年 4 月中旬, 其概率值即根据 1991 年 4 月中旬至 1993 年 4 月上旬两年资料求取的不同区间分布概率确定. 这样做的目的是避免回顾性震例研究中常见的震后资料包含于震前背景分析的缺陷, 也模拟了实际地震预报分析的进程。

为对比分析不同时间窗尺度对计算结果的影响, 我们分别计算了日均值、5 日均值、10 日均值 (5 日滑动)、15 日均值 (5 或 10 日滑动)、20 日均值 (10 日滑动) 和 30 日均值 (10 日滑动) 6 种时间窗的合成信息熵。

乌鲁木齐地区前兆观测台网在 80 年代后期才建设改造完善, 因此本文所用 17 项资料始

于 1990 年 1 月, 截止到 1996 年 6 月. 作为参考, 本文还给出 1982~1996 年的地下水 5 项资料 (水氡 2 项, 硫化氢 1 项, 甲烷 1 项, 水位 1 项) 的计算结果.

6 种时间窗的计算结果表明, 日均值、5 日均值的计算结果显示的异常不明显, 其它 4 种时间窗的计算结果较为一致.

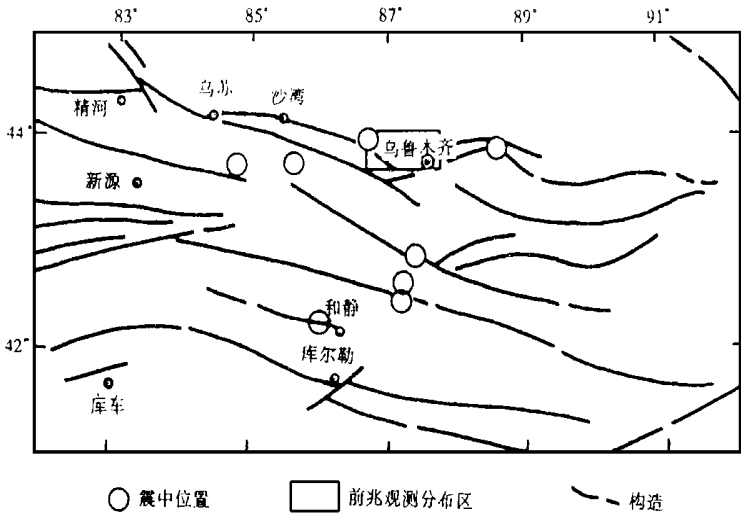


图 1 前兆观测区及震中分布图

Fig. 1 Distribution of precursor observation region and earthquakes.

图 2 为 1992~1996 年资料的 10 日均值 5 日滑动计算结果. 其间共发生 3 组地震. 在 1993 年 2 月 3 日和静 5.7 级地震和 1996 年 1 月 9 日沙湾 5.2 级地震前, 分别出现 1 年和 9 个月左右明显的低值异常减熵过程, 临震前 2 个月左右出现熵值回升. 在 1995 年 3 月 19 日和静 5.0 级地震和 1995 年 5 月 2 日乌苏 5.8 级地震前, 也存在熵值的低值异常变化过程, 但低值过程不如前两例明显.

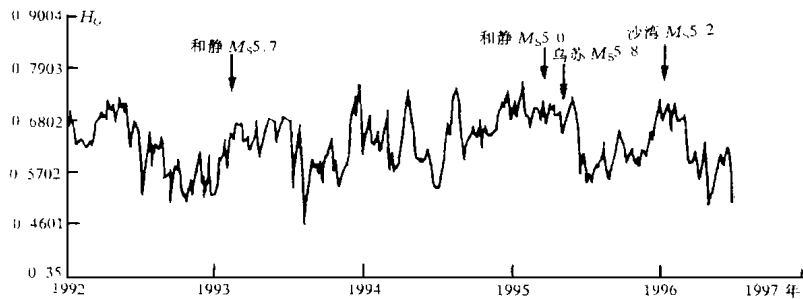


图 2 1992~1996 年多项资料 10 日均值 5 日滑动合成信息熵时间进程曲线

Fig. 2 Curve of ten-day mean values and five-day sliding synthetic information entropy of multi-precursor observing data(1992~1996).

图 3 为 1992~1996 年多项资料 20 日均值 10 日滑动计算结果. 异常曲线表明, 和静 5.7 级地震和沙湾 5.2 级地震前异常变化过程较为典型, 但在和静 5.0 级和乌苏 5.8 级地震前异常变化不太显著.

图 4 为 1982~1996 年地下水 5 项资料 10 日均值 5 日滑动计算结果. 由图可见, 在 1983

年 3 月呼图壁 5.4 级、1983 年 12 月和静 5.2 级、1991 年 6 月和静 5.2 级、1995 年 5 月乌苏 5.8 级及 1996 年 1 月沙湾 5.2 级等地震前, 分别出现半年至 1 年左右的不同程度的低熵过程. 另外, 1984 年下半年和 1988 年下半年~1989 年上半年出现了两个明显的低值异常过程, 但没有地震对应.

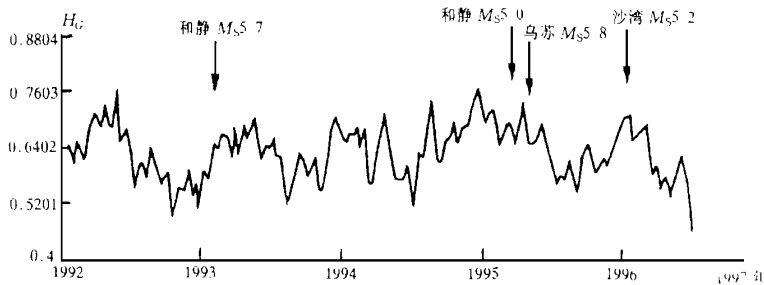


图 3 1992~1996 年多项资料 20 日均值 10 日滑动合成信息熵时间进程曲线

Fig. 3 Curve of twenty-day mean values and ten-day sliding synthetic information entropy of multi-precursor observing data(1992~1996).

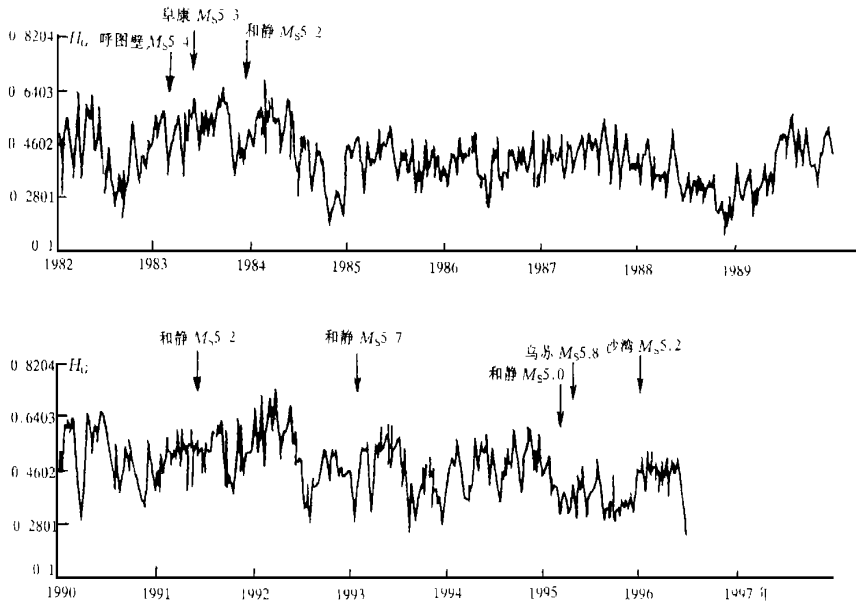


图 4 1982~1996 年地下水 5 项资料 10 日均值 5 日滑动合成信息熵时间进程曲线

Fig. 4 Curve of ten-day mean values and five day sliding synthetic information entropy of 5 items of groundwater data (1982~1996).

5 讨论和结语

熵是描述系统无序(或有序)程度的一个重要参数.对地震孕育过程而言,熵可以作为描述地震前兆复杂性的定量依据.震前的减熵异常变化在一定程度上显示了地震从孕育到发震阶

段, 系统经历了从无序向有序的演化过程. 这一点对于客观认识地震的孕育机制是有益的.

在实际震例的分析中, 发现震前的低熵变化过程多伴随有增熵起伏现象, 这是否是震源系统在接近临界状态前的“涨落”现象, 尚需进一步研究.

合成信息熵的计算避免了各单项观测资料的异常识别和综合分析中的人为性. 基于各种前兆观测出现异常变化是小概率事件的观点, 充分提取了观测资料中所携带的异常信息. 通过分析原始观测曲线, 发现各单项资料变化形态各异, 相互之间无明显的相关性, 可以认为所研究的前兆系统内各单项前兆之间是相互独立的. 因此根据信息合成原理, 将各种不同物理量的前兆观测项目的震兆观测异常信息综合起来, 可以达到信息增益的目的, 为综合判别异常提供较为客观的依据. 在通常的情况下, 只要是有一定物理基础的观测项目, 除了对资料进行一些必要的趋势或周期改正外, 无需做更复杂的数学处理就可以用于日常的会商预报工作. 相信通过对更多震例的深入研究, 合成信息熵会逐步成为一种量化的综合预报判据.

作为前兆资料合成信息熵计算的基础, 前兆资料的概率化处理是非常重要的, 它为前兆观测资料的无量纲化提供了一条新的途径.

参考文献

- 1 朱令人, 等. 澜沧-耿马 7.6 级地震前中小地震活动信息熵和分维的研究. 内陆地震, 1990, 4(4): 310~318.
- 2 周硕愚, 等. 用信息论方法研究大地震前的突变异常. 地震学报, 1986, 8(增刊): 501~512.
- 3 周硕愚, 等. 蕴震系统信息合成方法(ISSS). 中国地震, 1990, 6(4): 35~42.
- 4 簿万举, 等. 信息合成方法及其应用研究. 地壳形变与地震, 1995, 15(3): 84~88.
- 5 簿万举, 等. 形变、应变短临前兆信息的合成. 地震, 1995, (增刊): 148~157.
- 6 周硕愚. 系统科学导引. 北京: 地震出版社, 1988. 53~58.

STUDY ON SYNTHETIC INFORMATION ENTROPY OF MULTI-PRECURSOR OBSERVING DATA

WANG Haitao YANG Maling YANG Youling QU Yanjun

XU Qiulong ZHAO Cuiping ZHU Yan

(Seismological Bureau of Xinjiang Uygur Autonomous Region, Urumqi 830011)

Abstract

Temporal series of the multi-precursor observing data has been translated into probability temporal series by using the probabilization processing. Based on the work above, the synthetic information entropy of the information source system which is made up of the multi-precursor observing items has been calculated by applying the basic principle of information theory. Calculating the 17 items of precursor observation data in Urumqi and its neighborhood, it has been found that the anomaly process of entropy decrease appeared half a year to one year before moderate and strong earthquakes.

Key words Data processing Earthquake precursor, Probabilization, Synthetic information entropy, Group anomaly information